Reversibility and Irreversibility within the Quantum Formalism

Tim Jacobs and **Christian Maes**

Instituut voor Theoretische Fysica

K.U.Leuven, B-3001 Leuven, Belgium.

Setting

The discussion on time-reversal in quantum mechanics exists at least since Wigner's paper [17] in 1932. If and how the dynamics of the quantum world is time-reversible has been the subject of many controversies. Some have seen quantum mechanics as fundamentally time-irreversible, see for example von Neumann [16, p. 358], and some have seen in that the ultimate cause of time's arrow and second law behavior. In his best-selling book, Roger Penrose argues similarly and concludes that "*our sought-for quantum gravity must be a time-asymmetric theory*" [13, p. 351]. Not so long ago, we read yet about another project in Physicalia, [5]: to extend quantum mechanics into new fundamentally irreversible equations, thus proposing a new theory giving "*... une description fondamentale irréversible de tout système physique*".

We take here the opportunity to state and review a number of general points that are perhaps less emphasized in the existing literature. We first explain what is meant by mechanical reversibility and how it applies in classical mechanics and, somewhat differently, for the free evolution in quantum mechanics. If the evolution is not free, i.e., it is interrupted by measurements, the quantum formalism is challenged by the problem of retrodiction. Related to that is the notion of statistical reversibility which is very similar to what is more commonly known as the condition of detailed balance, at least for stochastic processes describing the spatio-temporal fluctuations in equilibrium. Finally, we describe the emergence of thermodynamic irreversibility. Here there is little difference between the classical and the quantum set-up even though one could have thought that the presence of a discrete energy spectrum prohibits dissipation for Hamiltonian evolutions.

Mechanical reversibility

Consider a classical mechanical system going through a sequence of positions and momenta $(q_0, p_0), (q_1, p_1) \ldots, (q_t, p_t)$. That evolution solves Newton's

equations of motion for given forces, e.g. gravity. Upon playing the movie backwards, i.e., time reversed, we see the system evolving from the positions q_t to q_0, but now with reversed momenta $-p_t$ to $-p_0$. The time-reversed sequence $(q_t, -p_t), \ldots, (q_0, -p_0)$ can or cannot be a solution of the same equation of motion. For say free fall, the time-reversed sequence certainly solves the same Newton's law of gravity; with friction or for the damped oscillator, that time-reversal symmetry is not present. In other words, symmetry with respect to time-reversal amounts to having identical mechanical laws for prediction and for retrodiction.

Generalizing to the more abstract idea of a dynamical system for which the state x_t at an arbitrary time t is given in terms of a flow $x_t = f_t(x_0)$ which is invertible, we say that the system is *mechanically* (or also, *dynamically*) *reversible* if there exists a transformation π of states with $\pi^2 = 1$, for which at all times

$$\pi f_t \pi = f_t^{-1} \tag{1}$$

The transformation π is often called the *kinematical* time-reversal. For classical Hamiltonian systems, f_t should be thought of as the Hamiltonian flow on the (micro)states $x = (q, p)$ given in terms of the positions q and the momenta p of all the particles, and π as the involution $\pi(q, p) = (q, -p)$. Mechanical reversibility expresses that first evolving forward in time and then changing all momenta gives the same state as first changing all momenta and then going back in time[1].

Statistical reversibility I

At many instances our understanding of physical phenomena involves statistical considerations[2]. We start by explaining what is a time-reversible

---

[1]Time-reversal in classical systems goes of course beyond the equations of Newton or Hamilton. A recent application of time-reversal in acoustics is found in [9].

[2]Even when God does not play dice and also for the description of the classical world, depending on the scale of the phenomena, stochastic dynamics enter. They can be the result of pure modelling or they appear as an effective or reduced dynamics. Classical examples are the Langevin description of Brownian motion, the Onsager-Machlup description of fluctuating hydrodynamics and the stochastically driven Navier-Stokes equation for turbulence. Since Pauli's work [12], stochastic processes are also obtained as the result of quantum considerations, for example also under situations where the Golden Rule applies. In fact, a great deal of so called classical stochastic processes, like Glauber and Kawasaki dynamics, find their origin in studies of quantum relaxation processes. Moreover, practically all quantum processes for open systems that are obtained under the weak coupling limit are just standard Markov processes.

stochastic process.

Restricting ourselves to stationary Markov processes, the law of the dynamics is given in terms of transition probabilities $p(x, s \to x', t)$ to find x' at time t, when the system is in state x at time s. To make it simple, we suppose discrete time and a finite state space. We thus have a sequence or trajectory $X_0, X_1, \ldots, X_t, \ldots$ of variables as sampled from a probability law that specifies

$$\text{Prob}[X_t = x] = \rho(x), \quad \text{the stationary distribution}$$

and the transition probabilities

$$\text{Prob}[X_t = x' | X_{t-1} = x, X_{t-2} = a_2, \ldots, X_0 = a_t] \equiv p(x, x')$$

We look at that stochastic dynamics in the time-window $[0, n]$ but running the movie backwards[3], i.e., in terms of $Y_t \equiv X_{n-t}$. Obviously, the time-reversed process is also Markovian and with the same stationary probability distribution ρ:

$$\text{Prob}[Y_t = y] = \text{Prob}[X_{n-t} = y] = \rho(y)$$

Its conditional probabilities are obtained (not: defined) from Bayes' rule

$$\text{Prob}[Y_t = y' | Y_{t-1} = y] = \text{Prob}[X_{n-t+1} = y | X_{n-t} = y'] \frac{\text{Prob}[X_{n-t} = y']}{\text{Prob}[X_{n-t+1} = y]}$$

$$= p(y', y) \frac{\rho(y')}{\rho(y)} \tag{2}$$

As a consequence, the reversed movie will be subject to the same statistical law whenever and only if the transition probabilities for the time-reversed process above

$$p(y', y) \frac{\rho(y')}{\rho(y)} = p(y, y')$$

equal that of the original process (right-hand side). That is certainly the case when the transition probabilities take the form

$$p(y, y') = \Phi(y, y') \, e^{[V(y) - V(y')]/2}, \quad \Phi(y, y') = \Phi(y', y) \tag{3}$$

for which $\rho(y) = e^{-V(y)}/Z$ is stationary[4]. The condition (3) is referred to as that of *detailed balance*.

---

[3]For simplicity we assume that the kinematical time-reversal π is identity, or that the variable X is even under time-reversal.

[4]Z is a normalization.

Such a scenario can be realized for a reduced description starting either from classical or from quantum mechanics. If we consider a classical Hamiltonian system $H = \frac{p^2}{2m} + V(q)$ on the constant energy surface, there is a natural and invariant measure: the so called Liouville volume-element measuring the phase space volume $|M|$ of a region M. Supposing of course that the energy is an even function of the momenta, $|\pi M| = |M|$ as there are as many states with positive or with negative momenta on the energy surface.

Let us now select two regions A and B in phase space. They could for example select all states with a particular density- or velocity-profile. With respect to the Liouville measure, we may ask what is the probability to find the state in the region B at time t when, at time $t = 0$, the state was in the region A. The formula is

$$\mathrm{Prob}[y_t \in B | y_0 \in A] = \frac{|A \cap f_t^{-1} B|}{|A|}$$

giving the volume-fraction of states in A that evolve to B. In the same way, replacing A with πB and B with πA,

$$\mathrm{Prob}[y_t \in \pi A | y_0 \in \pi B] = \frac{|\pi B \cap f_t^{-1} \pi A|}{|\pi B|} = \frac{|B \cap \pi f_t^{-1} \pi A|}{|B|}$$

Using (1), together with Liouville's theorem $|f_t^{-1} M| = |M|$, we conclude

$$\frac{\mathrm{Prob}[y_t \in B | y_0 \in A]}{\mathrm{Prob}[y_t \in \pi A | y_0 \in \pi B]} = e^{S(B) - S(A)} \tag{4}$$

which is the condition of detailed balance (3) where the Boltzmann (configurational) entropy $S(M) = \log |M|$ replaces the function V and ρ is being played by the Liouville measure. Indeed, when starting the classical mechanical system from statistical equilibrium, and one observes the statistical distribution of (some property) of the resulting trajectory, no distinction can be made between past and future. That is a direct consequence of (i) the mechanical reversibility and of (ii) the stationarity of the equilibrium. In particular, thermodynamic equilibrium can be characterized as the condition in which thermodynamic past and thermodynamic future are indistinguishable[5].

---

[5]And, in nonequilibrium situations, entropy production can be seen as a measure of the irreversibility, see [10].

Here comes a quantum example, at least in a toy-version. Let us consider just one quantum spin 1/2. We have the states *up* $|\uparrow\rangle$ and *down* $|\downarrow\rangle$ as basis. We suppose the following (Schrödinger) time-evolution

$$\psi_t = \frac{1}{2}\Big[\big(|\uparrow\rangle + |\downarrow\rangle\big) + e^{-i\omega t}\big(|\uparrow\rangle - |\downarrow\rangle\big)\Big] \tag{5}$$

where we have chosen to start at time $t = 0$ from $\psi_0 = |\uparrow\rangle$. At time $t = \pi/\omega$ we are in the *down* state $|\downarrow\rangle$ and at time $t = 2\pi/\omega$ we are back where we started. That is periodic motion with period $T = 2\pi/\omega$.

We measure the spin-magnetization, formally described in the observable

$$\sigma \equiv |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|$$

Simple algebra teaches that at time $t = T/4$ we find $\sigma = \pm 1$ with equal probability 1/2. At the same time and accordingly, the system is projected into the state $|\uparrow\rangle$ or into the state $|\downarrow\rangle$. Similarly, if we would have started at time $t = 0$ in the *down* state $|\downarrow\rangle$ and again measure at time $t = T/4$, exactly the same outcome statistics would occur as when started from the *up* state. Let us now imagine looking at the movie of outcomes when the spin is repeatedly measured at times $t = T/4, 2T/4, 3T/4, \ldots, nT/4, \ldots$. We see a random sequence $(\sigma_1, \sigma_2, \ldots, \sigma_n, \ldots)$ of outcomes $\sigma = +1$ or $\sigma = -1$ with stationary probability distribution $(1/2, 1/2)$. Obviously, when playing *that* movie backward, we are as bored as before. The statistics of the outcomes as seen in the time-reversed movie is identical to the original. Indeed, the process constituted by the successive measurement results is a time-reversible Markov process[6] (with trivial transition probabilities $p(\sigma, \sigma') = 1/2$).

The notion of spin in the above example is not at all to be taken serious. The above could as well describe the motion of a particle in a symmetric double well separated by a large barrier where in reality *up*, respectively *down*, stand for wave functions[7] supported (for the most part) in the *right* or in the *left* well; instead of measuring the spin we then speak about measuring the *right/left* position of the particle[8].

Note that everything above has relied heavily on the presence of an underlying stationary distribution. For the Markov process, it was the stationary

---

[6]In fact, it is a Bernoulli process.
[7]Say linear combinations of the ground state and the first excited state.
[8]The analogy is borrowed from Section 2.2 in [7].

distribution ρ that played an essential role in (2). For classical mechanics, it was the presence of the Liouville measure. In the quantum example, it was coin tossing.

Quantum free evolution

One of the most visible formal differences between the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \tag{6}$$

and Newton's law is that (6) is first order in time while Newton's $F = ma$ is second order[9]. Clearly then, the fact that $\psi(x,t)$ is a solution of (6) does not imply that $\psi(x,-t)$ is also a solution and in *that* sense Schrödinger's equation is not time-reversal invariant. We hasten to give the standard response, that one should also complex conjugate: $\psi(x,t)$ is a solution if and only if $\psi^\star(x,-t)$ is a solution. One argument comes from the representation of the momentum $p = -i\hbar\partial_x$, where the complex conjugation switches the sign of the momentum. One could reply to that by noting that there is no *a priori* reason that the momentum should change sign under time-reversal.[10] Furthermore it is not clear in general how to realize operationally a complex conjugation on the wave function of a system. Nevertheless, the more fruitful response is to complement time-reversal with a certain operation on wave functions much in the spirit of (1) as we will now explain.

One of the advantages of the abstraction around (1) is that the definition of mechanical reversibility also applies to the free evolution of the quantum formalism, i.e., the evolution on wave functions, say in the position-representation, as given by the standard Schrödinger equation (6). Following the proposal of Wigner [17], the recipe for time-reversal is to apply complex conjugation. More generally, the transformation π of above is now an anti-linear involution on Hilbert space. We get time-reversal symmetry when that π commutes with the quantum Hamiltonian H. Since the Schrödinger evolution is given by $U(t) \equiv e^{-itH/\hbar}$, equation (1) can now be written as

$$\pi U(t)\pi = U(t)^\dagger = U(t)^{-1} \tag{7}$$

---

[9]One could argue that Schrödinger's equation consists of two first order equations (since ψ is complex), very much analogous to Hamilton's equations of classical mechanics. We still think there is an essential difference but we do not wish to elaborate here on that issue as it is not essential for the rest of the paper.

[10]After all, in $p = -i\hbar\partial_x$, there is only a spatial (and no time-) derivative.

where still, in a way, $\pi(q, p) = (q, -p)$, albeit through a different mechanism.

We conclude that not only the (classical) Newton's law[11] but also the Schrödinger equation[12] are effectively invariant under dynamical time-reversal: for the free quantum flow, future and past are mere conventions and can be described by the same laws.

Quantum formalism

Since von Neumann [16], textbook quantum mechanics teaches us to complement the (linear) Schrödinger evolution by the so called reduction or collapse of the wave function to avoid the infamous measurement problem. The after-measurement wave function is obtained from the wave function before measurement by a highly nonlinear and *stochastic* transformation; the measurement is exactly the point where statistics enter the quantum formalism.

The role and status of the collapse and the associated measurement problem have been and still are extensively discussed in the literature[13]; that is not our main task here. Most physicists prefer not to speak about collapse of wave functions but give no alternative (or, what is worse, confuse decoherence with collapse). We hope that they would at least agree all the same that the collapse can be taken as an effective description of the entire measurement process. If, for the time being, we are happy with a pragmatic interpretation of quantum mechanics, then standard quantum mechanics works perfectly well and the measurement collapses the wave function *for all practical purposes*[14].

At this point, the plot thickens. Generally speaking, it is not possible to reverse the reduction. One cannot *de-measure* or *de-collapse* the wave function. The after-measurement wave function is very much limited — it must be an eigenstate of the measured observable — but not the before-measurement wave function. That is the point where for some problems and for others solutions seem to arise. Some find it odd that the rules of the game seem to break time-reversal symmetry on a rather fundamental level; it has less

---

[11]To be more precise, we should specify the forces — think of gravity. One can also include Maxwell's equations or Einstein's equations but that would take us too far.

[12]Or, for that matter, Dirac's equation. We do not wish to speak about possible time-symmetry breaking due to the weak interaction.

[13]... and at coffee-breaks or at lunch.

[14]We borrow the phrase from John Bell's paper [3].

esthetic appeal and nothing of it remains for the limiting classical mechanics. For others an opportunity seems created to give dynamical derivations of second law behavior.

Does that irreversible behavior of wave functions in the measurement formalism means that quantum mechanics is irreversible? Clearly the answer depends on how serious one takes that measurement formalism, or even on what one means by quantum mechanics. Remains that some see in it a manifestation of a fundamental spontaneous breaking of the time-symmetry at the beginnings of time[15].

Statistical reversibility II

The standard quantum theory with its measurement formalism is concerned exclusively with giving predictions for frequencies of future measurements. As a matter of logic, if one regards the collapse as a device that works for all practical purposes but has no ambition to be *fundamentally true*, there is no point to blame it for breaking mechanical time-reversal symmetry and to base major theoretical consequences on that. If one works on the level of a statistical theory, where one is happy to calculate probabilities of outcomes, to be consistent, one should only enquire about statistical reversibility. The time-reversal symmetry breaking of the collapse is then only pointing to a (theoretical) inadequacy of the standard interpretation with no further consequences except for giving yet another argument that the collapse procedure *cannot* be dynamically deduced from the (time-reversible and linear) Schrödinger evolution for the complete system + apparatus.

At first sight, the quantum spin example (5) seems a promising start to recover time-symmetry. Before we step back to meditate, we see whether we can generalize it to include for example higher dimensional projections[16].

On a more formal level, one considers a finite-dimensional Hilbert space and projections P_α where α runs over a set of linear subspaces. At the same time, α refers to an outcome of a measurement of some observable[17]. For example, we have a system of N distinguishable spin $\frac{1}{2}$ particles and we look at the

---

[15]We refer to the quantitative analysis of R. Penrose in Chapters 7 and 8 of [13].

[16]Other generalizations, like considering *fuzzy* measurements are of interest but will be skipped here.

[17]For notational simplicity we take in what follows measurements of the *same* observable. That is however not needed.

total magnetization in the z-direction

$$m_z \equiv \frac{1}{N} \sum_{i=1}^{N} \sigma_i^z \tag{8}$$

There are $N + 1$ outcomes $\{\alpha\}$ for a measurement of m_z. with corresponding orthogonal eigenspaces and projections P_α of different dimensionality $d_\alpha = \mathrm{Tr}[P_\alpha]$. The sum $\sum d_\alpha = d$ is the dimension of the Hilbert space. We refer to the α's as *conditions*. The time evolution on the spin-system is described by a Hamiltonian H, implemented by a unitary $U(t) = e^{-itH}$ ($\hbar$ equals *one*).

We start the system in condition α with probability d_α/d. That means that the initial density matrix is

$$\rho = \sum_{\{\alpha\}} \frac{d_\alpha}{d} \frac{P_\alpha}{d_\alpha} = 1/d$$

the normalized unit matrix. We measure the magnetization at fixed times[18] $t = 0, 1, 2, \ldots$ and we ask for the probability to find then the system consecutively in conditions $\alpha_0, \alpha_1, \alpha_2, \ldots$. Writing for short the sequence of outcomes $\omega = (\alpha_0, \alpha_1, \ldots, \alpha_t)$, that probability equals

$$\mathrm{Prob}[\omega] = \frac{1}{d} \mathrm{Tr}[U P_{\alpha_{t-1}} \ldots U P_{\alpha_1} U P_{\alpha_0} U^\dagger P_{\alpha_1} U^\dagger \ldots P_{\alpha_{t-1}} U^\dagger \ P_{\alpha_t}] \tag{9}$$

where $U \equiv U(t = 1) \equiv e^{-iH}$.

Again one can look at the time-reversed sequence. Like in the classical case, not only do we have to reverse the order of the conditions but we should also replace each projection P_α by its kinematical time-reversal $\pi P_\alpha \pi = P_{\alpha'}$. By that last procedure every condition α has a counterpart α'. We now ask about the probability to measure, in the indicated order, and starting from the same density matrix $\rho = 1/d$, the conditions

$$\Theta\omega = (\alpha_t', \alpha_{t-1}', \ldots, \alpha_0')$$

Using (7) that probability equals

$$\mathrm{Prob}[\Theta\omega] = \frac{1}{d} \mathrm{Tr}[U^\dagger P_{\alpha_1} \ldots U^\dagger P_{\alpha_{t-1}} U^\dagger P_{\alpha_t} U P_{\alpha_{t-1}} U \ldots P_{\alpha_1} U \ P_{\alpha_0}] \tag{10}$$

---

[18]Choosing unequal time-intervals between successive measurements makes no difference.

Upon inspection

$$\mathrm{Prob}[\omega] = \mathrm{Prob}[\Theta\omega]$$

In contrast with the quantum example (5), the statistics of the repeated measurements is in general no longer described by a Markov process but it does satisfy time-reversibility: from the statistics of outcomes we will not be able to decide whether the movie ran forward or backward. Note however again that we have here a stationary situation; at every moment the probability of condition α is d_α/d.

If instead of looking at the joint multi-time probability one considers the transition probabilities, one easily checks as done in [4] that a condition similar to detailed balance of (3) holds true. That is very much identical to what was obtained in (4) for the classical dynamics but with the Liouville volumes now being replaced by dimensions and the classical entropy replaced by $S(\alpha) = \log d_\alpha$, the quantum Boltzmann entropy associated to the condition α:

$$\frac{\mathrm{Prob}[\omega|\alpha_0]}{\mathrm{Prob}[\Theta\omega|\alpha_t']} = \frac{d_{\alpha_t}}{d_{\alpha_0}} \equiv e^{S(\alpha_t)-S(\alpha_0)} \tag{11}$$

One could wonder in the above (as in classical mechanics) what is the role of the stationary density matrix chosen to be a normalized multiple of unity. Moreover, that seems to prevent extensions to infinite dimensional Hilbert spaces. That problem can be avoided by considering the ensemble where one conditions on the results of the initial and final measurements. One then asks for the probability of outcomes $(\alpha_1, \alpha_2, \ldots, \alpha_{t-1})$ at consecutive times *given* the outcome α_0 at time 0 and the outcome α_t at time t *and given* that there were measurements at the intermediate times. That is

$$\mathrm{Prob}[\alpha_1 \ldots, \alpha_{t-1} \,|\, \alpha_0, \alpha_t] = \frac{\mathrm{Prob}[\omega]}{\sum_{\alpha_1,\ldots,\alpha_{t-1}} \mathrm{Prob}[\omega]} \tag{12}$$

Again it is easy to see that (12) is manifestly time-symmetric[19]. That observation was first made by Aharonov, Bergmann and Lebowitz in 1964 [2]. We repeat that the reversibility as in (12) is for the conditional probabilities of the results of a sequence of measurements, given the results of the initial and final measurements. That is entirely compatible with the irreversibility in the behavior of wave functions in the measurement formalism.

---

[19]Conditioning on past and future events is not so unphysical as one could imagine at first. In various experimental situations one selects the sample upon verifying both initial and final states.

Retrodiction

While we learn in school that prediction is difficult, especially of the future, and retrodiction is more easy, especially of the past, interesting questions extend in both directions. Textbooks in quantum mechanics usually emphasize the problem of prediction. Conventional quantum mechanics does not think of upcoming measurements as determining the state of the system now; the only information that is produced is about the state of the system subsequent to the execution of the measurement. In that sense the theory has a time-asymmetrical twist. Nevertheless the issue of retrodiction in quantum mechanics is not philosophical and it appears for example in problems of quantum optics and cryptography. From the point of view of textbook quantum mechanics, strange effects can appear. Here is one example.

Suppose a spin $1/2-$particle is initially prepared at time t_1 with spin pointing *up* in the $x-$direction. For time-evolution we take the trivial one, with Hamiltonian zero, so that the quantum state is unchanged through the Schrödinger evolution. At a later time $t_2 > t_1$ we measure its $z-$component to find it, suppose, *up* in the z-direction. Then we are sure that in the intermediate times $t \in (t_1, t_2)$ the spin was also *up* in the $z-$direction. By the latter we mean that, *if* we had measured the $z-$component, it would certainly have been *up*: a trivial dynamics would not be able to transform a *down* measurement into the *up* measurement at time t_2. This is an example of the time-symmetry when fixing initial and final results as explained at the end of the previous section, and as contained in [2]. Analogously, had we measured the $x-$component of the spin at these intermediate times, it would also certainly have been *up*: the trivial dynamics retains the initially prepared state! As a consequence, at an intermediate time, both the x- and the z-component seem well defined as their measurements would have been unambiguous. That statement is already somewhat strange for quantum mechanics, but it is very strange if you would think it could only be made *because* of the knowledge of the result of the later measurement, see also [1][20].

Let us now take two distinguishable spin $1/2$ particles, again with the trivial dynamics and initially prepared in some state

$$\psi_0 = c_{\uparrow\uparrow} |\uparrow\uparrow\rangle + c_{\uparrow\downarrow} |\uparrow\downarrow\rangle + c_{\downarrow\uparrow} |\downarrow\uparrow\rangle + c_{\downarrow\downarrow} |\downarrow\downarrow\rangle$$

At time t_1 we measure the total magnetization $\sigma_1^z + \sigma_2^z$ in the z-direction and

---

[20]Less trivial and more sensational aspects of retrodiction in quantum mechanics are discussed in [11, 14, 15].

we suppose that we find $\sigma_1^z + \sigma_2^z = 0$. At time $t_2 > t_1$ we measure $\sigma_1^z - \sigma_2^z$ and the probability that it is 2 (given the specified past) is equal to[21]

$$\frac{|c_{\uparrow\downarrow}|^2}{|c_{\uparrow\downarrow}|^2 + |c_{\downarrow\uparrow}|^2}$$

We now time-reverse. Assume that at time t_2 we in fact measure $\sigma_1^z - \sigma_2^z = 2$. That corresponds to the state $|\uparrow\downarrow\rangle$, which now serves as an initial state for the reversed dynamics. Then, under this reversed dynamics, certainly the magnetization $\sigma_1^z + \sigma_2^z = 0$ when measured at time $t_1 < t_2$.

Hence, the probability for measuring $\sigma_1^z - \sigma_2^z = 2$ at time t_2 given zero magnetization at time t_1 and started from ψ at time $t < t_1 < t_2$ depends on ψ through the coefficients c_{ij}, and can for example be equal to $\frac{1}{2}$. At the same time the probability to measure $\sigma_1^z + \sigma_2^z = 0$ at time t_1 before one sees $\sigma_1^z - \sigma_2^z = 2$ at time t_2 just equals one[22]!

All that is not in contradiction with the statistical reversibility of the previous section. It only shows that even on a microscopic level the dimensionality of the eigenspaces of the various observables involved in the art of retrodiction, plays a crucial role. Observe that when all c_{ij} are chosen equal, the ratio of forward to backward probabilities equals $1/2$, precisely the ratio of corresponding eigenspace dimensions as in the detailed balance condition (11). There are classical analogues of the example but not on the microscopic level; one must go to stochastic or reduced descriptions as for example in the understanding of thermodynamic irreversibility.

Thermodynamic irreversibility

Dynamical or statistical reversibility does not exclude macroscopic or thermodynamic irreversibility. That is relevant because it says we do not need explicit breaking of time-reversal invariance of the microscopic laws to account for the observed thermodynamic irreversibility. In fact, it is largely unimportant whether the microscopic time-evolution is reversible or not: macroscopic irreversibility typically obtains. Moreover, the mechanism that leads to second law behavior or entropy increase is quite independent of the classical or

---

[21] One should use the matrix $s^z = \frac{\hbar}{2}\sigma^z$ as the spin-observable. To ease the notation, we use σ^z instead.

[22] We think of that as the example on page 357 of Penrose's book [13] translated to spin-language.

quantum nature of the system.

Thermodynamic irreversibility is a statement about the *typical* temporal behavior of macroscopic observables. Something quite remarkable can happen: as the number N of degrees of freedom of the system gets very large and when the system is observed over the appropriate time scales, these macroscopic observables can start to follow an autonomous evolution. Usually one considers density or velocity profiles satisfying the kinetic or the hydrodynamic equations that constitute the phenomenology of time-dependent thermodynamics. Moreover, these equations are not only true *on average*; they are *typically* true, in the sense of the law of large numbers applied to the initial condition: the spreading or variance around the observable should go to zero for large N, while the expectation value solves a first order differential equation with a given initial value.

Obviously, to actually prove the emergence of autonomous macroscopic behavior with relaxation from far-from-equilibrium to equilibrium is mostly beyond our abilities today. However, that should not distract us from the physical mechanism behind thermodynamic irreversibility: as $N \to \infty$, the phase space volumes (classical systems) or the dimensions of the subspaces (quantum systems) corresponding to a particular macrostate, become overwhelmingly different. Conspiracies and special initial conditions aside, the dynamics brings the system to a region in phase space where the Boltzmann entropy is larger, precisely due to this huge difference in scales; a conclusion one draws purely by enumerating and counting the allowed possibilities. In the classical formalism, this set-up also explains why the so called Poincaré recurrences are irrelevant when studying relaxational behavior. In the quantum formalism, the same is true for the quasi-periodicity that occurs in finite Hamiltonian systems.

In sufficiently simple set-ups we do succeed in proving autonomous relaxation equations for the macroscopic variables and we thereby confirm the general picture above. An example is treated in full detail in [6] as a quantum extension of a model that was first conceived by Mark Kac [8, p. 99]. That quantum Kac model shows irreversible behavior for the magnetization as a function of time and the corresponding entropy in the model increases in time towards its equilibrium value. Yet the dynamics is given by standard quantum mechanics for a finite system of spins.

Conclusions

1. If one includes the collapse of the wave function in the mechanics of the quantum world, then there is no reversibility even on the microscopic level. One cannot demeasure things. One can however associate fundamental importance to that kind of microscopic irreversibility only up to the point where one considers the collapse as a fundamental input of quantum mechanics rather than as an effective device *for all practical purposes.*

2. If one understands the collapse of the wave function within the standard statistical interpretation of quantum mechanics it is appropriate to ask for statistical reversibility, i.e., in terms of probabilities of histories. In that case it is quite similar to the situation of *microscopic reversibility* or *detailed balance* for transition probabilities as obtained also classically from Hamiltonian dynamics. Statistical reversibility is satisfied within standard quantum mechanics.

3. Thermodynamic irreversibility is an emerging property in macroscopic behavior for which the reasons are basically unchanged in the transition from classical to quantum dynamics. In particular, such macroscopic irreversibility can be expected and sometimes is obtained on appropriate time-scales for quantum unitary evolutions with respect to typical initial data.

As a final remark, we like to add that a fully mechanically reversible version of quantum mechanics exists which is, for all we know, empirically equivalent with standard quantum mechanics: the Bohmian equations of motion[23]. However one does not necessarily need to resort to a modification of standard quantum practice; in a practical sense standard quantum theory can account for both the macroscopic irreversibility and statistical reversibility, as present in our daily experience.

---

[23]They consist of an equation for the wave function, nothing else than the Schrödinger equation, complemented with an equation for the positions of all particles. We refer to [1] for a discussion on the issue of retrodiction in Bohm's theory.

References

[1] Y. Aharonov, D. Albert, *The Issue of Retrodiction in Bohm's Theory*, in: *Quantum Implications*, B.J. Hiley and F.D. Peat (eds.), Routledge & Kegan Paul, New York, 224–226 (1987).

[2] Y. Aharonov, P.G. Bergmann, J.L. Lebowitz, *Time symmetry in the quantum process of measurement*, Phys. Rev. B **134**, 1410–1416 (1964).

[3] J.S. Bell, *Against measurement*, Physics World **8**, 33–40 (1990).

[4] I. Callens, W. De Roeck, T. Jacobs, C. Maes, K. Netočný, *Quantum entropy production as a measure for irreversibility*, Physica D **187**, 383–391 (2004).

[5] M. de Haan, *Vers l'historicité en théorie quantique*, Physicalia **26**, 29–31 (2004).

[6] W. De Roeck, T. Jacobs, C. Maes, K. Netočný, *An Extension of the Kac ring model*, J. Phys. A: Math Gen. **36**, 11547–11559 (2003).

[7] D. Dürr, S. Goldstein, N. Zanghi, *Quantum Chaos, Classical Randomness, and Bohmian Mechanics*, J. Stat. Phys. **68**, 259–270 (1992).

[8] M. Kac, *Probability and related topics in physical sciences* (chapter 14), Interscience Publishers Inc., New York (1959).

[9] P. Schewe, B. Stein, *Reversing time to catch snipers*, Physics News Update **687**, http://www.aip.org/pnu/2004/split/687-1.html

[10] C. Maes, K. Netočný, *Time-reversal and entropy*, J. Stat. Phys. **110**, 269–310 (2003).

[11] D. Mermin, *Limits to Quantum Mechanics as a Source of Magic Tricks: Retrodiction and the Bell-Kochen-Specker Theorem*, Phys. Rev. Lett. **74**, 831–834 (1995).

[12] W. Pauli, *Über das H-Theorem vom Anwachsen der Entropie vom Standpunkt der neueren Quantenmechanik*, in "Probleme der modernen Physik", Sommerfeld Festschrift, Leipzig, 30–45 (1928).

[13] R. Penrose, *The Emperor's New Mind*, Oxford University Press, Oxford (1989).

[14] S. Ben-Menahem, *Spin-measurement retrodiction*, Phys. Rev A **39**, 1621–1627 (1989).

[15] L. Vaidman, Y. Ahoronov, D.Z. Albert, *How to ascertain the values of σ_x, σ_y, and σ_z of a spin-1/2 particle*, Phys. Rev. Lett. **58**, 1385–1387 (1987).

[16] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (translated by R.T. Beyer), Princeton University Press, Princeton (1955).

[17] E. Wigner, *Über die Operation der Zeitumkehr in der Quantenmechanik*, Nachr. Akad. Ges. Wiss. Göttingen **31**, 546–559 (1932).